新生儿疼痛评估量表测量学性能的系统评价再评价

沈巧[1]，唐语蔓[1]，冷虹瑶[1]，雷若冰[2]，郑显兰[1*]

基金项目：国家自然科学基金项目（72074038）；2019 年重庆市技术创新与应用发展专项面上项目（cstc2019jscx-msxmX0157）

1. 400014 重庆 重庆医科大学附属儿童医院护理部 国家儿童健康与疾病临床医学研究中心 儿童发育疾病研究教育部重点实验室 儿科学重庆市重点实验室

2. 400014 重庆 重庆医科大学附属儿童医院儿童卓越证据与指南协同创新实验室

*通信作者：郑显兰，主任护师/教授，博士生导师；E-mail: zhengxianlan@vip.163.com

系统评价再评价注册号：CRD42021292583

【摘要】 **背景** 及时、准确评估新生儿疼痛是确保疼痛管理方案安全、有效实施的关键。目前，国内外学者针对不同的新生儿群体和疼痛类型，构建和测试了 40 多种疼痛评估量表。但是，其测量结果的可靠性和稳定性存在较大差异。并且，现有单项研究或系统评价仅能提供某个量表或量表某种测量性能的零散证据，将不利于医护人员的临床决策。**目的** 对新生儿疼痛评估量表测量学性能的系统评价进行再评价，为临床实践者和研究者选择最佳的疼痛评估量表提供证据支持。**方法** 计算机检索中英文数据库中近 5 年发表的新生儿疼痛评估量表测量学性能的系统评价。由两名研究者独立进行文献筛选和资料提取，并分别采用澳大利亚乔安娜布里格斯研究所(Joanna Briggs Institute, JBI)循证卫生保健中心对系统评价的方法学质量评价工具、系统评价偏倚风险（Risk of Bias in Systematic Review, ROBIS）工具和 Preferred Reporting Items for Systematic Reviews and Meta-Analyses（PRISMA）声明对纳入研究的方法学质量、偏倚风险和报告质量进行评价，进一步依据定性系统评价证据分级工具（Confidence of the Evidence from Reviews of Qualitative research，CERQual）评价纳入研究的证据质量。**结果** 共纳入 7 篇系统评价。文献质量评价结果显示，4 篇研究方法学质量较高，为低偏倚风险；3 篇研究方法学质量相对较低，为高偏倚风险。PRISMA 评价结果显示，5 篇研究报告相对完全（报告完成率＞60.00%），1 篇研究存在一定报告缺陷（报告完成率为 45.95%），1 篇研究有严重信息缺失（报告完成率为 10.81%）。CERQual 结果显示，新生儿疼痛评估量表测量性能结局的 22 条证据中，有高级证据 2 条（9.09%），中级证据 8 条（36.36%），低级证据 9 条（40.91%），极低级证据 3 条（13.64%）。证据综合结果显示，有 25 种量表用于新生儿疼痛评估时具有良好的内部一致性、评估者间信度、结构效度和可解释性，分别适用于早产儿和（或）足月儿的急性疼痛、持续性疼痛、术后疼痛或机械通气性疼痛。**结论** 尚无单个疼痛评估量表可用于新生儿所有类型疼痛的评估。推荐根据患儿年龄和疼痛类型等具体的临床情景选择经过验证的评估量表对患儿疼痛进行定期动态评估。并需进一步补充设计合理、方法严谨的高质量研究来检验现有量表测量我国新生儿人群各种类型疼痛时的可靠性和稳定性，探索扩宽其应用范围的可行性。

【关键词】新生儿；疼痛测定；量表；心理测定学；系统评价再评价

Psychometric properties of the pain assessment scales for neonatal infants: an overview of systematic review

SHEN Qiao[1], TANG Yuman[1], LENG Hongyao[1], LEI Ruobing[2], ZHENG Xianlan[1*]

1. Department of nursing, Children's Hospital of Chongqing Medical University, Chongqing 400014, China; National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Key Laboratory of Pediatrics, Chongqing, China.

2. Chevidence Lab Child & Adolescent Health, Department of Pediatric Research Institute, Children's Hospital of Chongqing Medical University, Chongqing, China

Corresponding authors: ZHENG Xianlan, Professor, Doctoral supervisor; E-mail: zhengxianlan@vip.163.com

PROSPERO registration number: CRD42021292583.

【Abstract】Background Timely and accurate assessment of neonatal pain are essential for pain management implemented safely and effectively. At present, scholars at home and abroad have constructed and tested more than 40 pain assessment scales for different neonatal groups and pain types. However, the reliability and stability of the outcomes varied. In addition, existing single studies or systematic reviews (SRs) can only provide scattered evidence of a certain scale or certain psychometric properties of the scale, which is not conducive to the clinical decision-making of medical staff. **Objective** To overview the systematic reviews of psychometric properties of the pain assessment scales for neonatal infants, and to provide evidence support for clinical practitioners and researchers for the best pain assessment scale. **Methods** Electronic searches were conducted to collect SRs related to psychometric properties of the pain assessment scales for neonatal infants published in the recent 5 years. Two researchers independently screened literature and extracted data, then assessed the methodological

quality, the reporting quality, and bias risk of the included studies separately with the JBI tools for SRs, ROBIS tools, and PRISMA statement, and further evaluated the evidence quality according to CERQual. **Results** A total of seven SRs were included in this study. According to the JBI tools and ROBIS tools, 4 studies were of high quality and were at low risk of bias and 3 studies were of relatively low quality and were at high risk of bias. With the PRISMA checklist, 5 SRs were found adequately reported over 60%, 1 SR was found to have certain reporting deficiencies with 45.95% of report completion rate, and 1 SR had serious information missing with 10.81% of report completion rate. The results of CERQual showed that there were 2 high-quality evidence (9.09%), 8 moderate-quality evidence (36.36%), 9 low-quality evidence (40.91%), and 3 critically-low-quality evidence (13.64%). A synthesis of evidence showed that there were 25 scales with good internal consistency, inter-rater reliability, construct validity, and interpretability, suitable for acute pain, persistent pain, postoperative pain, or mechanically ventilated pain respectively in preterm and/or term infants. **Conclusions** There is no single pain assessment scale available to assess all types of pain in neonates. It is recommended to select validated scales for neonatal pain combining the specific clinical context, such as the age group and type of pain. More well-designed and methodologically rigorous high-quality studies are needed to examine the reliability and stability of the existing scales in measuring various types of pain in neonates in China and to explore the feasibility of expanding their applications.

【Key words】 Newborn; Pain measurement; Scale; Psychometric evaluation; Overview

住院新生儿因诊疗原因，不可避免地要经历各种疼痛刺激。研究显示 82.6%的常见临床操作甚至可引发重度疼痛[1]。新生儿疼痛评估与镇痛管理专家共识指出，规范适宜的疼痛管理，可减轻或防止疼痛对身体或心理带来的各种不良影响，加速康复进程[2]。及时、准确评估新生儿疼痛是实施疼痛管理，衡量镇痛措施有效性的关键。由于新生儿无法通过语言表达疼痛，医护人员需借助神经生理、行为反应的测量对其疼痛进行评估。包含有生理和/或行为指标的疼痛评估量表被认为是最为便捷且费用最低廉的评估方法，医护人员经过简单培训即可掌握[3]。

基于共识选择健康测量工具的标准（Consensus-Based Standards for the selection of Health Measurement Instruments，COSMIN）指出只有具备良好测量学性能的量表才能作为可靠工具用于临床实践[4]。目前，针对不同的新生儿群体和不同的疼痛类型，已构建和测试了 40 多种疼痛评估量表。然而，对于具体的临床情景，应该选择哪种量表评估新生儿疼痛仍没有定论。其主要原因在于现有新生儿疼痛评估量表数量众多，且其测量结果的可靠性和稳定性存在较大差异[5]；而已发表的单项研究或系统评价仅能提供关于某个量表或量表某一方面测量学性能的零散证据，不利于医护人员的临床决策。系统评价再评价作为一种对某一特定问题相关系统评价进行综合研究的一种方法，可为证据使用者提供全面系统的高质量证据[6]。因此，本研究拟通过系统评价再评价，对现有新生儿疼痛评估量表的测量学性能进行全面评价，从而为临床实践者和研究者选择最佳的疼痛评估量表提供证据支持。

1 资料与方法

1.1 文献纳入和排除标准　纳入标准：①研究主题为新生儿疼痛评估量表的测量性能研究，包括信度（内部一致性信度、评分者间信度、重测信度等）、效度（内容效度、结构效度、校标效度等）、反应度，以及可解释性、实用性和可行性等（后三者虽并非测量性能，但是量表的重要特性）；②量表测评对象为住院新生儿；③研究类型为：系统评价、Meta 分析或 Meta 整合。排除标准：①研究主题为其他类型的疼痛评估方法，如新生儿疼痛表情自动识别系统、近红外光谱技术、心率变异性分析等；②无法提取测量学性能相关数据的文献；③发表语言为非中文或非英文；④系统评价计划书、会议摘要、重复发表的文献。

1.2 文献检索策略　2021 年 10 月 31 日将自由词和主题词结合制定检索策略（以中国知网和 PubMed 为例的中英文检索策略见框 1～2），计算机检索中国知网、中国生物医学文献数据库、万方、维普、PubMed、Embase、The Cochrane Library、Web of Science、CINAHL 数据库，并补充检索 Google 学术，通过纳入研究的参考文献追溯与本研究相关的文献。指南推荐意见的形成应基于当前可得的最佳证据[7]，而系统评价证据的有效期约为 3～5 年[8]，故，本次文献检索时段限制为近 5 年。

框 1 中国知网检索策略

Box 1 Search strategies for CNKI

（主题=疼痛 + 镇痛） AND （主题=新生儿 + 足月儿 + 早产儿 + 低出生体重儿 + 小于胎龄儿 + 婴儿） AND （主题= 评估 + 评分 + 测量 + 测定 + 量表 + 简表 + 问卷 + 工具） AND （主题=测量学 + 信效度 + 信度 + 效度 + 反应度 + 反应性 + 可行性 + 实用性 + 适用性 + 可操作性） AND （主题=系统评价 + 系统综述 + 荟萃分析 + Meta 分析 + Meta 整合） NOT （主题=剖宫产 + 分娩 + 产科 + 产妇 + 成人 + 老年人 + 妊娠）

框 2 Pubmed 检索策略

Box 2 Search strategies for Pubmed

#1 "pain"[Title/Abstract] OR "pain"[MeSH Terms]
#2 (newborn OR neonat* OR "newborn infant"[Title/Abstract] OR premature OR "preterm infant"[Title/Abstract] OR "premature infant"[Title/Abstract] OR "premature baby"[Title/Abstract] OR "premature babies"[Title/Abstract] OR "neonatal prematurity"[Title/Abstract] OR "term infant"[Title/Abstract] OR "low birth weight infant"[Title/Abstract] OR "low birth weight"[Title/Abstract] OR LBW OR ELBW OR VLBW OR "small for gestational age"[Title/Abstract]) OR ("infant, newborn" OR "Infant, premature" OR "infant, extremely premature" OR "infant, low birth weight" OR "infant, small for gestational age" OR "infant, very low birth weight"[MeSH Terms])
#3 (delivery[Title/Abstract] OR childbirth[Title/Abstract] OR labor[Title/Abstract]) OR ("Delivery, Obstetric" OR Parturition OR "Labor, Obstetric"[MeSH Terms])
#4 #1 AND #2 NOT (#3 AND #1)
#5 (measur*[Title/Abstract] OR assess*[Title/Abstract] OR evaluat*[Title/Abstract] OR rating[Title/Abstract] OR monitor*[Title/Abstract] OR quantif*[Title/Abstract] OR scale[Title/Abstract] OR tool[Title/Abstract] OR instrument[Title/Abstract] OR profile[Title/Abstract] OR questionnaire[Title/Abstract] OR score[Title/Abstract] OR intensit*[Title/Abstract] OR severit*[Title/Abstract]) OR ("pain measurement"[MeSH Terms])
#6 (psychometr*[Title/Abstract] OR reliability[Title/Abstract] OR consistency[Title/Abstract] OR accuracy[Title/Abstract] OR validity[Title/Abstract] OR responsiv*[Title/Abstract] OR sensitiv*[Title/Abstract] OR feasibility[Title/Abstract] OR utility[Title/Abstract] OR applicability[Title/Abstract]) OR ("reproducibility of results" OR "validation studies as topic" OR "feasibility studies"[MeSH Terms])
#7 #4 AND #5 AND #6, **Filters applied:** Meta-Analysis, Systematic Review, in the last 5 years

1.3 文献筛选与资料提取　利用 EndNote 软件对各数据库合并后的检索结果去除重复文献，进一步由 2 名研究者独立完成文献筛选和资料提取，并通过交叉核对与分歧讨论后达成一致意见。利用 Excel 提取资料，内容包括：作者、年份、标题、纳入研究数、研究对象、研究主题和主要结论等。

1.4 纳入研究的质量评价 由 2 名研究者分别独立地对纳入研究的方法学质量、偏倚风险、报告质量和证据质量进行评价，并交叉核对评价结果，采用组内讨论的形式解决分歧。具体的评价方法或标准如下：

1.4.1 文献质量评价　依据澳大利亚乔安娜布里格斯研究所(Joanna Briggs Institute, JBI)循证卫生保健中心对系统评价的方法学质量评价工具[9]对纳入研究的方法学质量进行评价。该系统评价的方法学质量评价工具包括循证问题、文献纳入排除标准、检索策略、文献质量评价、资料提取与综合、发表偏倚等 11 个条目。每个条目均可评价为“是”、“否”、“不清楚”和“不适用”。

依据系统评价偏倚风险（Risk of Bias in Systematic Review, ROBIS）工具[10]对纳入研究的偏倚风险进行评价。ROBIS 的评价过程分为三个阶段，阶段一：评估相关性（此部分主要针对干预性、病因性、诊断试验和预后性系统综述。故，本研究不进行阶段一的评估）；阶段二：评估系统评价制定过程中的偏倚风险程度，包括四个领域，共 21 个标志性问题；阶段三：评估系统评价整体的偏倚风险，包括三个标志性问题。评估领域及标志性问题（阶段二、三）的方法为：从系统评价中寻找支持偏倚风险程度判断的信息，回答标志性问题，判断偏倚风险程度。标志性问题的回答为“是”、“可能是”、“否”、“可能否”和“无信息”；偏倚风险程度的评价为“低（所有标志性问题的回答为“是”或“可能是”）”、“高（任一标志性问题的回答为“否”或“可能否”）”和“不确定”。

1.4.2 报告质量评价 由于测量工具系统评价的报告规范 PRISMA-COSMIN 正处于开发阶段，故本研究采用 Preferred Reporting Items for Systematic Reviews and Meta-Analyses（PRISMA）声明[11]对纳入研究的报告质量进行评价。PRISMA 声明作为系统评价报告的基础规范，包括 7 个部分，共 27 个条目。每个条目根据文献中是否有完整报告，

评价为“是”（完整报告）、“否”（部分报告+未报告）和“不适用”，并计算每个条目的报告完成率。

1.4.3 证据质量评价 纳入的系统评价由于原始研究间的临床异质性和方法学异质性均未进行定量合成。故，本研究主要依据定性系统评价证据分级工具 Confidence of the Evidence from Reviews of Qualitative research (CERQual)[12]对纳入研究证据质量进行评价：系统评价结果起始为高质量证据，然后从方法学局限性、相关性、结果一致性和数据充分性 4 各方面判断是否进行降级：不降级（无）、降 1 级（严重）、降 2 级（非常严重），最终将系统评价的证据质量分为“高”、“中”、“低”和“极低”四个级别。

2 结果

2.1 文献筛选流程及结果 初检共获得文献 122 篇，利用 Endnote 软件去除重复文献 34 篇，阅读标题和摘要剔除研究对象、研究主题、研究类型和发表语言等明显不符合纳入标准的文献 76 篇，进一步阅读全文剔除会议摘要 2 篇，研究对象不符 3 篇，最终纳入 7 篇文献进行描述性分析。文献筛选流程图见图 1。

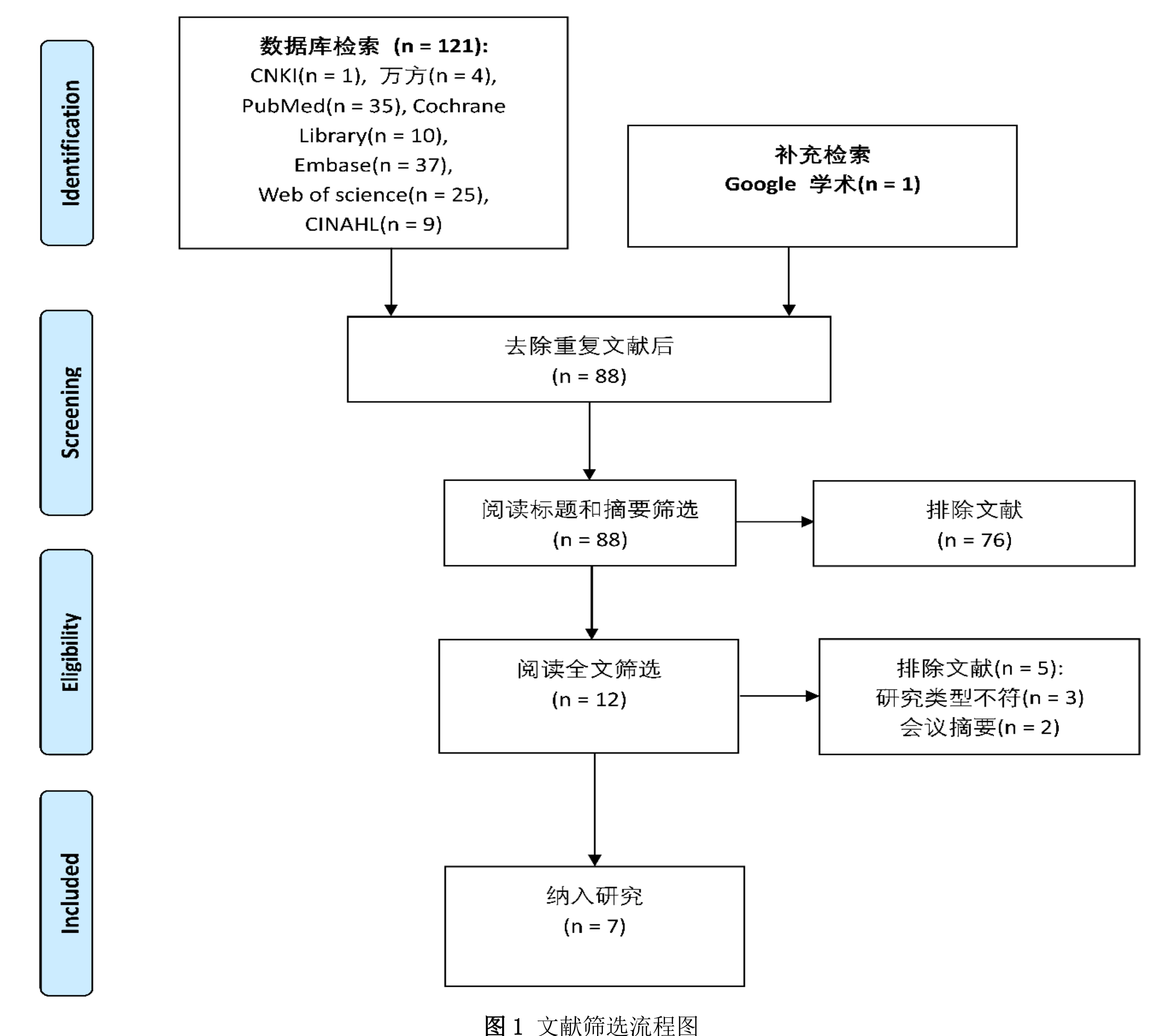


图 1 文献筛选流程图

Figure 1 Flowchart of the selection of included systematic reviews

2.2 纳入系统评价的基本特征 系统评价纳入研究的数量在 9～352 篇不等，涉及 1～40 种新生儿疼痛评估量表，可归纳为急性疼痛、持续性疼痛、术后疼痛、机械通气性疼痛四种类型。纳入的 1 篇发表在 JAMA Pediatrics 上的系统评价[13]对现有新生儿疼痛评估量表的测量学特征进行了综合性评价，其余纳入研究仅针对一个疼痛评估量表、一类新生儿群体，或一种测量学特征：2 篇系统评价[14, 15]分别评价了新生儿疼痛、躁动和镇静评（Neonatal Pain, Agitation, and Sedation Scale；N-PASS）和舒适量表（COMFORT）的测量学性能；2 篇系统评价[16, 17]分别关注了机械通气新生儿和颌面外科新生儿两个特殊人群；2 篇系统评价[18, 19]分别对新生儿疼痛评估量表的反应度和实用性进行了评价。纳入系统评价的基本特征见表 1。

2.3 文献质量评价结果 纳入研究文献质量评价结果显示，4 篇研究方法学质量较高，为低偏倚风险[13, 15, 18, 19]，3 篇研究方法学质量相对较低，为高偏倚风险[14, 16, 17]。影响纳入研究方法学质量及偏倚风险的主要因素为：研究纳入排除标准不适合系统评价的问题，或未清晰明确地界定纳入研究的特征；未提供完整的检索策略，或未包括重要的检索数据库，未通过除数据库以外的其他方法确定相关研究；研究采用的文献质量评价工具不符合系统评价的问题和纳入研究的文献类型，或未提取和报告量表测量性能的关键结局指标。纳入系统评价的方法学质量及偏倚风险评价结果分别见表 2 和表 3。

2.4 报告质量评价结果 PRISMA 评价结果显示 5 篇研究报告相对完全（报告完成率＞60.00%）[13-15, 18, 19]，1 篇研究存在一定报告缺陷（报告完成率为 45.95%）[16]；1 篇研究有有严重信息缺失（报告完成率为 10.81%）[17]。报告质量不足主要表现为结构式摘要，注册与计划书，检索策略，每个结果合成中纳入研究的处理过程及方法，研究结果间异质性分析、证据可信度评价，资金支持，数据、代码和其他材料的可用性报告不全面（规范报告率＜50%）。纳入系统评价的报告质量评价结果见表 4。

2.5 纳入研究的证据质量评价结果 采用 CERQual 方法评价新生儿疼痛评估量表的信度、效度和反应度三种测量性能结局，共 22 条证据的质量。结果显示，提取的证据中有高级证据 2 条（9.09%），中级证据 8 条（36.36%），低级证据 9 条（40.91%），极低级证据 3 条（13.64%）。证据质量降级的主要原因是原始研究的方法学局限性、研究结果间的异质性、纳入研究与系统评价问题的间接性，以及数据不充分性（小样本量研究）。由于 Shapour 等[17]的研究未标引纳入研究并报告研究特征，故未能获取相关信息评价其测量性能的证据质量。其余纳入系统评价的证据质量评价结果见表 5。

表 1 纳入系统评价的基本特征

Table 1 Characteristics of the included systematic reviews

纳入研究	研究主题	纳入研究数	研究对象	纳入量表数	量表使用情景	方法学质量评价	结果综合方法
Emma，2021[19]	疼痛评估量表实用性（使用频率及使用正确性）评价	352	新生儿	22	急性操作性疼痛 持续性疼痛 术后疼痛	仅评价了盲法	描述性统计分析
Hanna，2020[16]	疼痛评估量表的综合评价（信效度/医护认知）	12	机械通气新生儿	8	机械通气性合并急性操作性疼痛 单纯机械通气性疼痛	Newcastle-Ottawa Scale（NOS）	描述性分析
Mikayla，2020[14]	新生儿疼痛、躁动和镇静评分 N-PASS 的信效度评价*	29	新生儿	1	急性疼痛 持续性疼痛	牛津循证医学中心质量评价标准	描述性分析
Vito，2019[13]	疼痛/镇静评估量表在不同人群和不同临床情景应用时的信效度、反应度、区分度、适用性等的综合评价	89	足月儿 早产儿 幼儿	40#	急性疼痛 持续性疼痛 术后疼痛 机械通气性疼痛	COSMIN 偏倚风险评价清单*	描述性分析
Naomi，2019[18]	疼痛评估量表反应度评价	9	新生儿	10	术后疼痛 持续性疼痛	COSMIN 偏倚风险评价清单*	描述性分析
Maaskant，2016[15]	COMFORT 系列量表信效度和反应度评价	30	新生儿 0～18 岁儿童	4#	操作性疼痛 术后疼痛	COSMIN 偏倚风险评价清单*	描述性分析
Shapour，2016[17]	疼痛评估量表的信效度和实用性（使用频率）评价	17	颌面外科新生儿	16	术后疼痛	—	描述性分析

* N-PASS：Neonatal Pain，Agitation，and Sedation Scale/新生儿疼痛、躁动和镇静评；COSMIN：Consensus-Based Standards for the selection of Health Measurement Instruments/基于共识选择健康测量工具的标准；

研究中纳入的新生儿疼痛评估量表的数量

表 2 纳入系统评价的方法学质量评价结果

Table 2 Methodological quality of the included systematic reviews

纳入系统评价	条目 1	条目 2	条目 3	条目 4	条目 5	条目 6	条目 7	条目 8	条目 9	条目 10	条目 11
Emma 等 2021[19]	是	是	是	是	否	是	是	是	不适用	是	是
Hanna 等 2020[16]	是	否	不清楚	否	否	不清楚	是	是	不适用	是	不清楚
Mikayla 等 2020[14]	是	否	不清楚	是	否	是	是	是	不适用	是	是
Vito 等 2019[13]	是	是	不清楚	是	是	是	不清楚	是	不适用	是	是
Naomi 等 2019[18]	是	是	是	是	是	不清楚	不清楚	是	不适用	是	是
Maaskant 等 2016[15]	是	是	不清楚	是	是	是	是	是	不适用	是	否
Shapour 等 2016[17]	否	是	不清楚	否	不清楚	不清楚	不清楚	不清楚	不适用	是	否

备注：1.所提出的循证问题是否清晰明确？2.文献的纳入标准是否恰当？3.采用的检索策略是否恰当？4.研究论文的来源是否恰当？5.采用的文献质量评价标准是否恰当？6. 是否由 2 名或 2 名以上的评价者独立完成文献质量评价？7.提取资料时是否采用一定的措施减少误差？8.综合/合并研究的方法是否恰当？9.是否对可能的发表偏倚进行评估？10.作者是否在报道数据的支持下对政策和（或）实践提出推荐意见？11.对今后进一步研究的特定方向是否提出恰当建议？

表 3 纳入系统评价的偏倚风险评价结果

Table 3 Risk of bias in the included systematic reviews

纳入系统评价	阶段二				阶段三
	领域 1：研究的纳入排除标准	领域 2：研究的检索和筛选	领域3：数据提取和质量评价	领域4：数据合成和结果呈现	系统评价的偏倚风险
Emma 等 2021[19]	低	低	高	低	低
Hanna 等 2020[16]	高	高	高	低	高
Mikayla 等 2020[14]	高	低	高	低	高
Vito 等 2019[13]	低	低	低	低	低
Naomi 等 2019[18]	低	低	低	低	低
Maaskant 等 2016[15]	低	低	低	低	低
Shapour 等 2016[17]	低	高	不确定	高	高

表 4 纳入系统评价的报告质量评价结果

Table 4 Reporting quality the included systematic reviews

文章结构	PRISMA 条目	Emma (2021)	Hanna (2020)	Mikayl (2020)	Vito (2019)	Naomi (2019)	Maaskant (2016)	Shapour (2016)	规范报告率
标题	1 标题	是	是	是	是	是	是	是	100.00%
摘要	2 结构式摘要	否	是	否	是	否	否	否	28.57%
背景	3 理论基础	是	是	是	是	是	是	是	100.00%
	4 目的	是	是	是	是	是	是	否	85.71%
方法	5 纳排标准	是	是	是	是	是	是	是	100.00%
	6 信息来源	是	否	是	是	是	是	否	71.43%
	7 检索策略	是	否	否	否	是	否	否	28.57%
	8 研究选择	是	是	否	是	是	是	否	71.43%
	9 资料提取	是	是	是	否	否	是	否	57.14%
	10a 资料条目-结局指标数据	是	否	是	否	是	是	否	57.14%
	10b 资料条目-其他变量	是	是	是	是	是	是	否	85.71%
	11 偏倚风险评价	是	否	是	是	否	是	否	57.14%
	12 效应指标	不适用	不适用	不适用	不适用	不适用	不适用	不适用	—
	13a 方法综合-结果合并时纳入研究的过程	是	否	否	否	否	否	否	14.29%
	13b 方法综合-数据准备	是	否	否	否	是	否	否	28.57%
	13c 方法综合-单个研究/综合结果呈现方法	是	否	否	否	否	否	否	14.29%
	13d 方法综合-结果综合方法	是	否	是	是	否	是	否	57.14%
	13e 方法综合-异质性分析方法	否	否	是	是	否	是	否	42.86%
	13f 方法综合-敏感性分析方法	不适用	不适用	不适用	不适用	不适用	不适用	不适用	—
	14 报告偏倚评价	不适用	不适用	不适用	不适用	不适用	不适用	不适用	—
	15 可信度评价	否	否	否	否	否	否	否	0.00%
结果	16a 研究选择-检索和筛选过程结果	是	是	是	是	是	是	否	85.71%
	16b 研究选择-排除研究	是	是	是	是	是	是	否	85.71%
	17 研究特征	是	是	是	否	是	是	否	71.43%
	18 研究偏倚风险	是	是	是	是	是	是	否	85.71%
	19 单个研究结果	是	否	是	否	是	是	否	57.14%
	20a 结果综合-纳入研究特征及偏倚风险	是	否	是	是	是	是	否	71.43%
	20b 结果综合-统计综合结果	是	否	是	是	是	是	否	71.43%
	20c 结果综合-异质性原因	否	否	是	是	否	是	否	42.86%

	20d 结果综合-敏感性分析结果	不适用	不适用	不适用	不适用	不适用	不适用	不适用	—
	21 报告偏倚	不适用	不适用	不适用	不适用	不适用	不适用	不适用	—
	22 证据可信度	否	否	否	否	否	否	否	0.00%
讨论	23a 讨论-结果解释	是	是	是	是	是	是	否	85.71%
	23b 讨论-证据的局限性	是	是	是	是	是	是	否	85.71%
	23c 讨论-系统评价过程局限性	是	是	是	是	是	否	否	71.43%
	23d 讨论-结果对实践等的影响	是	是	是	是	是	是	是	100.00%
其他信息	24a 注册与计划书-注册信息	是	否	否	是	否	否	否	28.57%
	24b 注册与计划书-计划书	是	否	否	是	否	否	否	28.57%
	24c 注册与计划书-注册/计划书的信息修改	否	否	否	否	否	否	否	0.00%
	25 支持	是	否	否	否	是	是	否	42.86%
	26 利益冲突	是	是	是	是	是	是	否	85.71%
	27 数据、代码和其他材料的可用性	是	否	否	是	是	否	否	42.86%
报告完成率		83.78%	45.95%	64.86%	67.57%	64.86%	67.57%	10.81%	

表 5 纳入系统评价的证据质量评价结果

Table 5 Evidence quality the included systematic reviews

纳入系统评价	纳入研究数	研究结果	方法学局限性	相关性	结果一致性	数据充分性	CERQual证据质量
Hanna 等 2020[16]	5	COMFORTneo 和 N-PASS 量表用于测量机械通气新生儿疼痛反应的信效度较好	无	无	严重[a]	无	中
Mikayla 等 2020[14]	4	N-PASS 评估新生儿急性疼痛的信度：N-PASS 测量机械通气或非机械通气状态下所有胎龄（足月儿/早产儿）新生儿急性疼痛的可靠性较好：内部一致性信度范围为 α = 0.837-0.971;评估者间的信度为 ICC= 0.93-0.99	严重[b]	无	无	无	中
	3	N-PASS 评估新生儿急性疼痛的效度：通过与 PIPP 和 FLACC 比较，评估 N-PASS 用于测量机械通气或非机械通气状态下所有胎龄新生儿急性疼痛的结构效度：Spearman 秩相关系数范围从中等正相关到高度正相关（ρ = 0.62-0.75）。N-PASS 与 FLACC 相关性高（r= 0.980-0.996）	严重[b]	无	严重[a]	无	低
	1	N-PASS 评估新生儿持续性疼痛的信度：	无	无	无	严重[C]	中

		N-PASS 测量机械通气、非机械通气或术后状态下所有胎龄（足月儿/早产儿）新生儿持续性疼痛的可靠性较好：内部一致性信度范围为 α = 0.72-0.82，评估者间的信度为 r=0.97					
	3	N-PASS 评估新生儿持续性疼痛的效度：通过与 PIPP 比较评估 N-PASS 测量机械通气或术后状态下新生儿持续性疼痛的结构效度：Spearman 秩相关系数范围从中等到高度正相关（ρ = 0.62-0.83）	无	无	严重[a]	无	中
Vito 等 2019[13]	12	NFCS、N-PASS、COMFORT 和 COMFORT-B 可用于评估早产儿和足月儿的急性疼痛和持续性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	无[d]	严重[e]	无	无	中
	1	PIPP 可用于评估早产儿和足月儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	无[d]	无	无	无	高
	1	COMFORTneo 可用于评估早产儿和足月儿的持续性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	无[d]	无	无	无	高
	1	EVENDOL 可用于评估足月儿的急性疼痛和持续性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	无[d]	严重[e]	无	无	中
	1	EDIN 可用于评估早产儿的持续性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	无[d]	严重[e]	无	严重[C]	中
	3	MAPS、CHIPPS 可用于评估早产儿和足月儿的术后疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[f]	严重[e]	无	无	低
	2	PAT 可用于评估早产儿和足月儿的术后疼痛和持续性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[f]	严重[e]	无	无	低
	1	BIIP 可用于评估早产儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[f]	严重[e]	无	严重[C]	极低
	1	MBPS 可用于评估足月儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[f]	严重[e]	无	严重[C]	极低
	4	ABC 和 COVERS 可用于评估早产儿和足月儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[f]	严重[e]	无	无	低
	5	APN/DAN、NIPS 和 NIAPAS 可用于评估早产儿和足月儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[g]	严重[e]	无	无	低
	6	FLACC、POPS、RIPS 和 NAPI 可用于评估足月儿的术后疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[g]	严重[e]	无	无	低
	1	UWCH 可用于评估足月儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[g]	严重[e]	无	严重[C]	极低

	1	PASPI 可用于评估早产儿的急性疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[f]	无	无	严重[c]	低
	2	POCIS 可用于评估足月儿的急性疼痛、持续性疼痛和术后疼痛，具有较好的内部一致性、评估者间信度、结构效度和可解释性	严重[g]	严重[e]	无	无	低
Naomi 等 2019[18]	9	仅 10 个新生儿疼痛评估量表进行了反应性研究，其中 8 个量表的疼痛评分在镇痛干预后显著降低，反应度较好：7 个新生儿疼痛评估量表 (N-PASS、COMFORT-B、EVENDOL、COMFORTneo、FLACC、EDIN、MAPS)干预前后评分差异较大，1 个量表(LIDS)干预前后评分差异中等	严重[b]	无	无	无	中
Maaskant 等 2016[15]	5	COMFORT 量表/COMFORT-without blood pressure 用于评估新生儿操作性疼痛，COMFORTneo 用于评估持续性疼痛时的内部一致性信度和评估者间信度均＞0.70，信度较好；结构效度的相关系数在 0.38 和 0.74 之间	无	无	严重[a]	严重[c]	低

备注：a. 研究间信效度结果差异较大；b.纳入的大多数研究的质量较差；c.小样本量研究；d. COSMIN 偏倚风险评价均为低偏倚；e.研究对象包括幼儿；f.COSMIN 偏倚风险评价均为中等偏倚；g.COSMIN 偏倚风险评价均为高偏倚

APN/DAN: Acute Pain in Newborns/Douleur Aiguëdu Nouveau-né/新生儿急性疼痛行为评分量表；**BIIP**: Behavioral Indicators of Infant Pain/婴儿疼痛行为指征量表；**CERQual**：Confidence of the Evidence from Reviews of Qualitative research/定性系统评价证据分级工具； **CHIPPS**：Children’s and Infants’ Postoperative Pain Scale/小儿术后疼痛评分；**EDIN**: Échelle Douleur Inconfort Nouveau-Né/新生儿疼痛与不适量表；**FLACC**：Face, Legs,Activity, Cry and Consolability/儿童疼痛行为量表；**ICC**: Interclass Correlation Coefficient/组内相关系数；**LIDS**: Liverpool Infant Distress Score/利物浦婴儿不适评分；**MAPS**：Multidimensional Assessment of Pain Scale/多维疼痛评估量表；**MBPS**：Modified Behavioral Pain Scale/改良行为疼痛量表；**NAPI**: Nursing Assessment of Pain Intensity/疼痛程度护理评估单；**NFCS**: Neonatal Facial Coding System/新生儿面部表情编码系统； **NIPS**: Neonatal Infant Pain Scale/新生儿疼痛量表；**NIAPAS**：Neonatal Acute Pain Assessment Scale/新生儿急性疼痛评估量表；**N-PASS**: Neonatal Pain, Agitation and Sedation Scale/新生儿疼痛、躁动及镇静评估量表；**PAT**: Pain Assessment Tool/疼痛评估工具；**PASPI**: Pain assessment scale for preterm infants/早产儿疼痛评估量表；**PIPP**: Premature Infant Pain Profile/早产儿疼痛量表；**POCIS**: Pain observation scale for young children/幼儿疼痛观察量表；**POPS**: Postoperative Pain Score/术后疼痛评分；**RIPS**: Riley Infant Pain Scale/ Riley 婴儿疼痛量表；UWCH：University of Wisconsin Children’s Hospital Pain Scale/威斯康星大学儿童医院疼痛量表。

3 讨论

3.1 新生儿疼痛评估量表的选择与推荐 测量性能是选择最佳患者报告结局测量工具（Patient reported mutcome measures，PROMs）首要考虑的因素之一[20]。一个经过信效度和反应度验证的评估量表可准确反应新生儿的疼痛状态，指导医护人员及时调整镇痛方案，达到最好的疼痛预防和控制效果。本研究采用系统评价再评价的方法对新生儿疼痛评估量表测量性能相关证据进行了严格评价与综合。纳入研究共评价了四十余种新生儿疼痛评估量表的测量性能。证据综合结果（表 5）显示，有 25 种量表用于新生儿疼痛评估时具有良好的内部一致性、评估者间信度、结构效度和可解释性。其中 8 种量表 COSMIN 偏倚风险评价为低风险、6 种量表为中风险、5 种量表为高风险，分别适用于早产儿和（或）足月儿的急性疼痛、持续性疼痛、术后疼痛或机械通气性疼痛；尚无足够证据支持单个疼痛评估量表用于所有新生儿各种类型疼痛的评估。

然而，现有疼痛量表的多样性及适用范围的局限性极大降低了国内外新生儿疼痛评估率，制约了疼痛管理的开展[21]。目前，国内外新生儿病房至少需同时使用 3～5 种量表评估患儿疼痛，这对量表使用者和医院管理者均是极大的挑战[22]。另外，由于国内外学者尚未对新生儿疼痛的分类及定义进行清楚界定，这将直接影响医护人员正确判断新生儿经历的疼痛类型，进而影响恰当疼痛评估量表的选择。所以，新生儿疼痛评估量表相关推荐意见的形成需要考虑：在平衡测量学性能的前提下，优先选择同时适用于多种疼痛类型的普适量表，以免医护人员对量表的选择不当导致低估或高估患儿疼痛程度。本研究结果显示，目前可同时适用于 2 种及以上疼痛类型评估的量表包括：新生儿面部表情编码系统（Neonatal Facial Coding System，NFCS）、N-PASS、COMFORT、COMFORT-B，可同时用于早产儿和足月儿急性疼痛和持续性疼痛（包括机械通气性疼痛）的评估（中级质量证据）；疼痛评估工具（Pain Assessment Tool，PAT）可用于早产儿和足月儿术后疼痛和持续性疼痛的评估（低级质量证据）；幼儿疼痛观察量表（Pain observation scale for young children，POCIS）可用于足月儿急性疼痛、持续性疼痛和术后疼痛的评估（低级质量证据）。

除测量性能和适用情景，临床实践者在选择恰当的评估工具时还需考虑 PROMs 的语言及临床实用性。由于以上量表均来源于国外，为保证量表本土化应用过程的语义对等性和内容可理解性，应采用标准流程对量表进行汉化[23]。然而，目前仅 N-PASS[24]和 COMFORT-B[25]有中文译本，且后者由于未进行回译，其汉化本测量概念的可靠性和有效性需进一步研究确认。N-PASS 经何碧云等引入翻译形成中文修订版[24]，并在急性操作性疼痛合并机械通气性疼痛[26,27]，单纯机械通气性疼痛[25]和单纯术后疼痛[28]三种疼痛类型中进行了验证，结果均显示 N-PASS 具有较好的信效度和反应度。临床实用性方面，Emma 等[19]的系统评价结果显示现有新生儿镇痛有效性的随机对照试验中有 2.80%的研究采用了 N-PASS 量表。国内调查研究显示，全国新生儿病房中有 4.31%采用 N-PASS 进行日常疼痛评估[21]。该量表亦得到了美国儿科学会（American Academy of Pediatrics，AAP）和中国医师协会新生儿科医师分会的推荐[2,29]。因此，鉴于 N-PASS 可能具备准确测量住院新生儿各种疼痛类型的潜力，建议国内研究者进一步扩大中文版 N-PASS 的验证范围，但需要基于规范的量表测量性能评价方法。同时，在更加证据出现之前，推荐根据患儿年龄和疼痛类型等具体的临床情景选择经过验证的评估量表对患儿疼痛进行定期动态评估[30]。并需进一步补充设计合理、方法严谨的高质量研究来检验现有量表测量我国新生儿人群各种类型疼痛时的可靠性和稳定性，探索扩宽其应用范围的可行性。

3.2 纳入系统评价的总体质量欠佳，疼痛评估量表系统评价研究的制作和报告过程仍需进一步规范 WHO 临床实践指南制定手册指出指南制定者需审慎评价系统评价质量，只有最新的高质量系统评价才能作为指南推荐意见的证据来源[7]。为选择最佳的 PROMs，制作高质量的系统评价至关重要[31]。为此，由多位心理测量学相关领域的专家组成的 COSMIN 工作组提

出了基于共识选择健康测量工具的标准 COSMIN，用于指导 PROMs 系统评价的制作以帮助临床实践者和研究者选择最恰当的健康测量工具[32]。然而，本研究中纳入的新生儿疼痛评估量表的系统评价并未遵照 COSMIN 系统评价的标准步骤，而主要遵循常规系统评价制作和报告的方法，但由于未能有效处理信效度和反应度的结局，致使无法对量表推荐意见及强度进行判断。此外，纳入研究中同样出现了现有系统评价或 meta 分析存在的共性问题[33]：①循证问题、纳排标准及检索策略的结构不清晰：未包含所测构念、目标人群、类型和测量属性四个关键要素[20]，甚至简单套用干预性研究循证问题的 PICO 模式[16]；②检索资源不全面：未包括 COSMIN 指南建议的两个基本数据库 Medline 和 Embase[31]，或未进行相关研究的补充检索；③文献质量评价标准不恰当：纳入系统评价中仅 3 篇研究[13, 15, 18]采用 COSMIN 偏倚风险清单对量表测量性能评价研究的方法学质量进行评价；④资料提取不充分：缺少量表测量性能的关键结局指标数据。⑤重要信息报告缺失：在摘要（规范报告率为 28.57%），方法（规范报告率为 48.98%），结果（规范报告率为 44.44%）和其他信息（规范报告率为 38.10%）四部分均存在重要信息缺失。以上研究及报告局限性将直接影响量表的证据质量和推荐等级。基于此，建议研究者参考 COSMIN 指南[31]对最佳测量性能评价研究的界定，严格设计、实施和报告量表测量性能评价研究；进一步采用基于共识选择健康测量工具的标准 COSMIN [32]规范制作新生儿疼痛评估量表测量学性能的系统评价，并形成最终推荐意见，以指导临床实践者和研究者选择最佳的疼痛评估量表。

3.3 本研究的局限性　本研究存在一定的局限性：研究结果主要基于当前发表的系统评价，其稳定性可能受到新发表文献或文献更新的影响；由于纳入研究的临床异质性和方法学异质性未能进行 meta 分析，进而未能定量比较各评估量表测量学性能的优劣；由于缺乏量表实用性的直接证据，研究将疼痛评估量表的使用频率作为实用性检验的依据可能不能反应量表实用性的真实情况。

4 小结　虽然当前证据尚不足以支持单个疼痛评估量表用于新生儿所有类型疼痛的评估，但目前有 25 种量表用于新生儿疼痛评估时具有良好的测量性能，分别适用于早产儿和（或）足月儿的急性疼痛、持续性疼痛、术后疼痛或机械通气性疼痛。建议临床实践者及研究者根据新生儿的胎龄和疼痛类型等具体的临床情景选择经过验证的量表进行患儿疼痛评估。同时，建议采用基于共识选择健康测量工具的标准 COSMIN 规范制作新生儿疼痛评估量表测量性能的评价研究，以验证现有量表测量我国新生儿人群各种类型疼痛时的可靠性和稳定性，并进一步制作系统评价，形成最佳疼痛评估量表的推荐意见。

作者贡献：沈巧为主要研究人员，负责此项研究的设计、实施、分析和报告。郑显兰作为论文的通讯作者和指导专家，负责研究实施的质量控制。沈巧、唐语蔓、冷虹瑶和雷若冰负责文献筛选，资料提取和质量评价。沈巧起草了文章初稿，所有作者都对稿件进行了严格的审阅和修改。

本文无利益冲突。

参考文献

[1] 王亚静, 李杨, 孙静, 等. 新生儿重症监护病房患儿操作性疼痛现状调查[J]. 护理学杂志, 2019, 34(11): 20-23.

[2] 程锐, 杨洋, 史源, 等. 新生儿疼痛评估与镇痛管理专家共识（2020 版）[J]. 中国当代儿科杂志, 2020, 22(09): 923-930.

[3] 万丽, 赵晴, 陈军, 等. 疼痛评估量表应用的中国专家共识(2020 版)[J]. 中华疼痛学杂志, 2020, 16(03): 177-187.

[4] 陈祎婷,彭健,沈蓝君,胡雁,余桂星,李铮.COSMIN 方法介绍 :结合一项系统评价实例进行解读[J].护理研究,2021,35(09):1505-1510.

[5] Maxwell LG, Fraga MV, Malavolta CP. Assessment of Pain in the Newborn: An Update[J].Clinics in perinatology, 2019, 46(4): 693-707.
[6] Higgins JPT, Thomas J, Chandler J(editors), et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.
[7] World Health Organization. WHO handbook for guideline development, 2nd ed[M]. Geneva: World Health Organization, 2014. https://apps.who.int/iris/handle/10665/145714.
[8] Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated?[J].Jama, 2001, 286(12): 1461-1467.
[9] 顾莺, 张慧文, 周英凤, 等. JBI 循证卫生保健中心关于不同类型研究的质量评价工具——系统评价的方法学质量评价[J]. 护士进修杂志, 2018, 33(08): 701-703.
[10] Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed[J]. J Clin Epidemiol. 2016, 69:225-34.
[11] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews[J]. Rev Esp Cardiol (Engl Ed). 2021,74(9):790-799.
[12] Lewin S, Booth A, Glenton C, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series[J]. Implement Sci. 2018, 13(Suppl 1):2.
[13] Giordano V, Edobor J, Deindl P, et al. Pain and Sedation Scales for Neonatal and Pediatric Patients in a Preverbal Stage of Development: A Systematic Review[J].JAMA Pediatr, 2019, 173(12): 1186-1197.
[14] Morgan ME, Kukora S, Nemshak M, et al. Neonatal Pain, Agitation, and Sedation Scale's use, reliability, and validity: a systematic review[J].Journal of perinatology : official journal of the California Perinatal Association, 2020, 40(12): 1753-1763.
[15] Maaskant J, Raymakers - Janssen P, Veldhoen E, et al. The clinimetric properties of the COMFORT scale: A systematic review[J].European Journal of Pain, 2016, 20(10): 1587-1611.
[16] Popowicz H, Kwiecień-Jaguś K, Olszewska J, et al. Pain Scales in Neonates Receiving Mechanical Ventilation in Neonatal Intensive Care Units - Systematic Review[J].J Pain Res, 2020, 13: 1883-1897.
[17] Yaripoor S, Khalili A, Joonbakhsh F, et al. Systematic Review of Pain assessment scales in newborns under maxillofacial surgery Admitted to the surgical ward[J].International Journal of Medical Research and Health Sciences, 2016, 5: 41-44.
[18] Meesters N, Dilles T, Simons S, et al. Do Pain Measurement Instruments Detect the Effect of Pain-Reducing Interventions in Neonates? A Systematic Review on Responsiveness[J].The journal of pain, 2019, 20(7): 760-770.
[19] Olsson E, Ahl H, Bengtsson K, et al. The use and reporting of neonatal pain scales: a systematic review of randomized trials[J].Pain, 2021, 162(2): 353-360.
[20] 陈祎婷,彭健,沈蓝君,等. COSMIN 方法介绍:制作患者报告结局测量工具的系统评价[J].护士进修杂志,2021,36(8):699-703.
[21] 沈巧, 郑显兰, 林紫, 等. 66 家医疗机构儿童疼痛管理现状调查[J]. 中国护理管理, 2019, 19(02): 187-193.
[22] Andersen RD, Munsters JMA, Vederhus BJ, et al. Pain assessment practices in Swedish and Norwegian neonatal care units[J]. Scand J Caring Sci. 2018, 32(3):1074-1082.
[23] 魏雪, 杨红. 医学量表的跨文化翻译：经验及启示[J].海外英语, 2021, 11: 207-208+231.

[24] 何碧云. 新生儿疼痛、躁动及镇静评估量表的汉化及信效度评价[D]. 浙江大学, 2017.
[25] 贾勇刚. 三种行为观察量表在评估烧伤婴幼儿患者疼痛中的应用研究[D]. 暨南大学, 2012.
[26] 庞玮, 张海林, 张小宝, 等. 不同量表用于机械通气新生儿急性疼痛评估的价值分析[J]. 现代生物医学进展, 2019, 19(02): 262-265+283.
[27] 黄小知. 三种量表用于机械通气新生儿急性疼痛评估的比较[D]. 南方医科大学, 2017.
[28] 何星蓉, 郑显兰, 柯淞淋, 等. 基于 3 个英文版新生儿疼痛量表评估新生儿术后疼痛的信度和效度研究[J]. 中国循证儿科杂志, 2021, 16(03): 186-191.
[29] COMMITTEE ON FETUS AND NEWBORN and SECTION ON ANESTHESIOLOGY AND PAIN MEDICINE. Prevention and Management of Procedural Pain in the Neonate: An Update[J].Pediatrics, 2016, 137(2): e20154271.
[30] 沈巧, 郑显兰, 李霞, 等. 儿童疼痛管理相关临床实践指南内容分析[J]. 护理学杂志, 2018, 33(07): 50-53.
[31] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures[J]. Quality of life research, 2018, 27(5): 1147-1157.
[32] Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs)–user manual [EB/OL]. http://www.cosmin.nl/, 2018-02/2022-03-11.
[33] 王志稳, 冷敏敏, 赵雅洁, 等. 近 5 年我国护理领域干预性研究系统评价的方法学质量分析[J]. 中国护理管理, 2020, 20(04): 485-489.